METHOD AND APPARATUS FOR ESTIMATING CAMERA MOTION

Field of the invention

The present invention relates to a method and an apparatus for estimating a camera motion; and, more particularly, to a method and an apparatus for estimating a camera motion, in which the motion of a camera made at the time of photographing can be extracted from a moving picture photographed by the camera to be used for moving picture classification and search.

Description of the Prior Art

PCT Appln. No. WO 00/42771 discloses "Camera Motion Parameters Estimation Method".

The disclosed camera motion estimation method includes the steps of extracting from a video sequence vectors corresponding to the motion between two successive frames, the motion vectors forming the camera velocity field; preprocessing the camera velocity field so as to reduce the amount of data and the heterogeneousness of the extracted motion vectors; estimating for each pair of frames, from the preprocessed field, camera features between the two considered frames; and undertaking a long term motion analysis based on

the estimation in order to obtain motion descriptors corresponding to the estimated camera motion parameters.

The camera motion estimation method constructed as described above is applied to a sequence of successive video frames divided into blocks, and can be employed to implement descriptors within Moving Picture Experts Group (MPEG) 7.

In the estimation method, the extraction of the motion vector is carried out using block matching motion compensation, which is employed as a part of the predictive coding process widely used in video transmission for reducing the amount of information needed to encode a video sequence. To be specific, each frame is divided into a fixed number of blocks, and for each block, a search is made for the most similar block in a previous reference frame, over a predetermined area. The search criterion is generally the search of the best matching block, giving the least prediction error. A motion vector is computed on the basis of the positions of each pair of similar blocks, and a camera parameter between each pair of frames is computed using the computed motion vector and a certain camera motion equation.

As described above, in the conventional camera motion parameter estimation method, a motion vector is extracted by comparing a reference frame with a current frame and a camera motion parameter is estimated using the motion vector. However, the conventional camera motion parameter estimation

method has drawbacks in that the speed of parameter estimation is not desirable.

Accordingly, the development of a camera motion parameter estimation method or a camera motion estimation apparatus, which is capable of increasing the speed of estimation, is highly desired.

Summary of the Invention

It is, therefore, an object of the present invention to provide an apparatus for estimating a camera motion, in which horizontal and vertical mean pictures are generated with respect to each of the blocks of each picture frame throughout the entire picture frames and motion vectors are extracted from the horizontal and vertical mean pictures, thereby increasing the speed of camera motion estimation.

Another object of the present invention is to provide a method for estimating a camera motion, in which a picture frame is divided into a plurality of blocks, horizontal and vertical mean pictures are generated with respect to each block, motion vectors are extracted from the horizontal and vertical mean pictures, and then the camera motion is estimated by the combination of the motion vectors.

In accordance with an aspect of the present invention, there is provided a camera motion estimation apparatus for

extracting a camera motion made at the time of photographing from a photographed moving picture, comprising: a frame picture memory unit for storing a plurality of moving picture frames with one frame separated from another; a buffer for sequentially extracting the frames stored in the frame picture memory unit and storing them; a block division unit for dividing each of the frames stored in the buffer into a plurality of blocks; a mean picture generation unit for generating horizontal and vertical mean pictures using the blocks obtained through the block division unit to estimate a camera motion; a horizontal and a vertical mean picture memory unit for storing the data of the horizontal and the vertical mean pictures generated in the mean picture generation unit; a motion vector extraction unit for extracting motion vectors from the blocks using horizontal and vertical mean picture data stored in the horizontal and the vertical mean picture memory unit; a motion vector memory unit for storing the data of the motion vectors extracted in the motion vector extraction unit; and a camera motion estimation unit for estimating a camera motion using the motion vector data stored in the motion vector memory unit.

In accordance with another aspect of the present invention, there is provided a camera motion estimation method of extracting a camera motion made at the time of photographing from a photographed moving picture, comprising

the steps of: dividing each of successive frames into a plurality of blocks and generating horizontal and vertical mean pictures with respect to each of the blocks; extracting a representative motion vector with respect to each of the blocks using the horizontal and the vertical mean pictures; estimating a partial camera motion within each of the blocks using the representative motion vector; and estimating the entire camera motion within the moving picture on the basis of the partial camera motions.

Brief Description of the Drawings

The above and other objects, features and other advantages of the present invention will be more apparently understood from the following detailed description taken in conjunction with the accompanying drawings, in which:

Fig. 1 is a block diagram of an apparatus for estimating a camera motion in accordance with the present invention;

Fig. 2 is a view for showing a divided picture frame;

Fig. 3 is a view for showing the generation of horizontal and vertical mean pictures; and

Fig. 4 is a view showing the motion vector extraction unit of the camera motion estimation apparatus according to the present invention.

Detailed Description of the Preferred Embodiments

As shown in Fig. 1, an apparatus for estimating a camera motion in accordance with the present invention includes a frame picture memory unit 20 for storing a plurality of moving picture frames 10 with one frame separated from another, a buffer 21 for sequentially extracting the frames stored in the frame picture memory unit 20 and storing them, a block division unit 22 for dividing each of the frames stored in the buffer 21 into a plurality of blocks, a mean picture generation unit 31 for generating horizontal and vertical mean pictures using the blocks obtained in the block division unit 22 to estimate a camera motion, a horizontal and a vertical mean picture memory unit 32 and 33 for storing data of the horizontal and vertical mean pictures generated in the mean picture generation unit 31, a motion vector extraction unit 41 for extracting motion vectors from the blocks using horizontal and vertical mean picture data stored in the horizontal and the vertical mean picture memory unit 32 and 33, a motion vector memory unit 42 for storing the data of the motion vectors extracted in the motion vector extraction unit 41, and a camera motion estimation unit 51 for estimating a camera motion using the motion vector data stored in the motion vector memory unit 42.

The motion estimation process that is performed by the

camera motion estimation apparatus is described hereinafter.

A plurality of moving picture frames 10 are sequentially stored in the frame picture memory unit 20.

When the moving picture frames stored in the frame picture memory unit 20 are color picture frames, the color picture frames are converted into gray scale picture frames and stored in the buffer 21.

The gray scale picture frames stored in the buffer 21 are each divided into a plurality of blocks having a uniform size by the block division unit 22. Each of the blocks is m in width and n in length. The magnitudes of the width and length should be smaller than those of each frame, respectively. Generally, the block has a size of $m = n = 8$.

The gray scale picture frames divided into the subblocks by the block division unit 22 are sequentially transmitted to the mean picture generation unit 31. As shown in Fig. 2, each of the gray scale picture frames is divided into the blocks B_1 to B_n, and the blocks B_1 to B_n are transmitted to the mean picture generation unit 31.

The mean picture generation unit 31 generates horizontal and vertical mean pictures using each of the transmitted blocks.

The horizontal and vertical mean pictures are generated as shown in Fig. 3. With reference to Fig. 3, a mean picture generation method is described.

Each block has 8 pixels in width and 8 pixels in length. B_d^k denotes the d-th block of a k-th frame, and H_n^k denotes the n-th horizontal line of the d-th block of a k-th frame.

The total mean value of a horizontal line H_n^k is mapped to the point (k,n) of a horizontal mean picture P_d^h of each block, while the total mean value of a vertical line V_n^k is mapped to the point (k,n) of a vertical mean picture P_d^v of each block. Such a mapping is carried out according to equation 1.

$$P_d^h(k,l) = \frac{1}{m}\sum_{i=1}^{m} B_d^k(i,l) \qquad (1)$$
$$P_d^v(k,l) = \frac{1}{n}\sum_{j=1}^{n} B_d^l(k,j)$$

The horizontal mean picture generated in the mean picture generation unit 31 is stored in the horizontal mean picture memory unit 32. Similarly, the vertical mean picture is stored in the vertical mean picture memory unit 33.

When processing of a frame is completed, the mean picture generation unit 31 transmits a request signal for the next frame. The frame picture memory unit 20 receives and stores the next frame in response to the request signal for the next

frame. The above-described process is repeated until processing of the final frame is completed. In this case, the horizontal and the vertical mean picture memory unit 32 and 33 hold the data of the previous pictures until processing of the final frame is completed.

The motion vector extraction unit 41 extracts a motion vector with respect to each block using the horizontal and vertical mean pictures stored in the horizontal and vertical mean memory unit 32 and 33. The motion vectors extracted are stored in the motion vector memory unit 42. When extraction of motion vectors with respect to all the frames is completed, the motion vector extraction unit 41 transmits a motion vector extraction completion signal to the camera motion estimation unit 51. At this time, the camera motion estimation unit 51 estimates a camera motion using motion vectors with respect to the frames stored in the motion vector memory unit 42.

The motion vector extraction unit 41, as depicted in Fig. 4, includes a block picture discrimination unit 411, an edge extraction unit 412, and a block motion vector extraction unit 413.

The block picture discrimination unit 411 converts data transmitted from the horizontal and the vertical mean picture memory unit 32 and 33 into block pictures.

The edge extraction unit 412 extracts edge components from the block pictures using the block pictures converted in

the block picture discrimination unit 411. The extracted edge components include a magnitude component and a direction component.

The block motion vector extraction unit 413 extracts the motion vector of each block picture from the edge components extracted through the edge extraction unit 412. A number of edge components can be extracted by the edge extraction unit 412. The block motion vector extraction unit 413 compares a plurality of edge components with one another and assigns a representative value for the motion vector of the block picture.

In the comparison of edge components, a plurality of edge components are compared with one another, and edge components having a magnitude component equal to or smaller than a certain reference value are removed. The reference value can be determined by a user. Then, the direction components of the remaining edge components are arranged and a direction component of the highest frequency is assigned for a representative direction component. Then, the motion vector extraction unit 41 stores the representative component in the motion vector memory unit 42.

The camera motion estimation unit 51 extracts a camera motion using the motion vectors extracted in the motion vector extraction unit 41. In detail, the camera motion estimation unit 51 receives the representative direction components of

the blocks, and estimates the camera motion using the representative direction components. In this case, the representative components used as input values are temporal direction components due to the characteristics of block pictures.

Spatial direction components are required to extract the camera motion, so that the representative direction components are converted into spatial direction components. When T is the total number of frames, this conversion is carried out according to Equation 2.

$$x_d = T/\tan\theta_{v,d} \qquad (2)$$

$$y_d = T * \tan\theta_{h,d}$$

The spatial movement speeds u_x and u_y are obtained using the movement values x_d and y_d in x and y -axis directions obtained through Equation 2 and the center position values x_o and y_o of the blocks. Where the position values of each block are x_o and y_o, $u_x = x_0 - x_d$ and $u_y = y_0 - y_d$. The camera motions R_x, R_y and R_{zoom} are calculated by substituting the spatial movement speeds u_x and u_y and the center position

values x_o and y_o for u_x, u_y, x and y of Equation 3, respectively.

$$u_x = \frac{xy}{f} R_x - f(1 + \frac{x^2}{f^2}) R_y + f \tan^{-1}(\frac{x}{f})(1 + \frac{x^2}{f^2}) R_{zoom} \qquad (3)$$

$$u_y = -\frac{xy}{f} R_y - f(1 + \frac{x^2}{f^2}) R_x + f \tan^{-1}(\frac{y}{f})(1 + \frac{y^2}{f^2}) R_{zoom}$$

wherein f is set to 1.

Camera motion parameters R_x, R_y and R_{zoom} obtained for each block through Equation 3 are compared with one another, and then the parameter of the greatest value is assigned for the representative camera motion of the block to thereby assign a representative camera motion of the highest frequency for the camera motion of a current moving picture.

As described above, the camera motion made at the time when a moving picture was photographed can be extracted from the moving picture. The extracted camera motion is utilized as a parameter for moving picture classification and search.

As described above, the present invention provides a method for estimating a camera motion, in which a picture frame is divided into a plurality of blocks, horizontal and vertical mean pictures are generated with respect to each block, motion vectors are extracted from the horizontal and vertical mean pictures, and then the camera motion is estimated by the combination of the motion vectors.

While the present invention has been shown and described with respect to the preferred embodiments, it will be understood by those skilled in the art that various changes and modifications may be made without departing from the spirit and scope of the present invention as defined in the following claims.